

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

(A) NAME: FORBAIRT (trading as BioResearch Ireland)
(B) STREET: Glasnevin
(C) CITY: Dublin 9
(E) COUNTRY: Ireland
(F) POSTAL CODE (ZIP): none

(A) NAME: UNIVERSITY COLLEGE CORK
(B) STREET: College Road
(C) CITY: Cork
(E) COUNTRY: Ireland
(F) POSTAL CODE (ZIP): none

(A) NAME: MCCARTHY, Thomas Valentine
(B) STREET: Vista Villa, Montenotte
(C) CITY: Cork
(E) COUNTRY: Ireland
(F) POSTAL CODE (ZIP): none

(A) NAME: VAUGHAN, Patrick Martin
(B) STREET: 175 West Avenue Parkgate, Frankfield
(C) CITY: Cork
(E) COUNTRY: Ireland
(F) POSTAL CODE (ZIP): none

(ii) TITLE OF INVENTION: A method for the characterisation of nucleic acid molecules involving generation of extendible upstream DNA fragments resulting from the cleavage of nucleic acid at an abasic site

(iii) NUMBER OF SEQUENCES: 32

(iv) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 93 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

000201" 6E2E2950

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(F) TISSUE TYPE: Skeletal muscle

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

TCCAAGGAGA AGCTGGATGT GGCCCCCAAG CGGGATGTGG AGGGCATGGG CCCCCCTGAG	60
ATCAAGTACG GGGAGTCACT GTGCTTCGTG CAG	93

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 93 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "DNA generated by PCR amplification"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

TCCAAGGAGA AGCTGGATGT GGCCCCCAAG CGGGAUGUGG AGGGCAUGGG CCCCCCUGAG	60
AUCAAGUACG GGGAGUCACU GUGCUUCGUG CAG	93

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 93 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "DNA generated by PCR amplification"

(iii) HYPOTHETICAL: NO

000201" 6E/E/960

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 93 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "DNA generated by glycosylase mediated cleavage followed by extension of upstream fragment"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

CTGCACGAAG CACAGTGACT CCCCGTACTT GATCTCAGGG GGGCCCATGC CCTCCACATC	60
CCGCTTGGGG GCCACATCCA GCTTCTCCTT GGA	93

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 273 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(F) TISSUE TYPE: Skeletal muscle

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

TCCAAGGAGA AGCTGGATGT GGCCCCAAG CGGGATGTGG AGGGCATGGG CCCCCCTGAG	60
ATCAAGTACG GGGAGTCACT GTGCTTCGTG CAGCATGTGG CCTCAGGACT GTGGCTCACC	120
TATGCCGCTC CAGACCCCAA GGCCCTGCGG CTCGGCGTGC TCAAGAAGAA GGCCATGCTG	180
CACCAGGAGG GCCACATGGA CGACGCACTG TCGCTGACCC GCTGCCAGCA GGAGGAGTCC	240
CAGGCCGCCC GCATGATCCA CAGCACCAAT GGC	273

000201" 6EZE960

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 273 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

(F) TISSUE TYPE: Skeletal muscle

TCCAAGGAGA	AGCTGGATGT	GGCCCCCAAG	CGGGATGTGG	AGGGCATGGG	CCCCCTGAG	60
ATCAAGTACA	GGGAGTCACT	GTGCTTCGTG	CAGCATGTGG	CCTCAGGACT	GTGGCTCACC	120
TATGCCGCTC	CAGACCCCAA	GGCCCTGCGG	CTCGGCGTGC	TCAAGAAGAA	GGCCATGCTG	180
CACCAGGAGG	GCCACATGGA	CGACGCACTG	TCGCTGACCC	GCTGCCAGCA	GGAGGAGTCC	240
CAGGCCGCCC	GCATGATCCA	CAGCACCAAT	GGC			273

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 196 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(A) DESCRIPTION: /desc = "DNA generated by glycosylase mediated cleavage and upstream fragment extension, and has a 3' hydrogen atom"

(iv) ANTI-SENSE: NO

(A) NAME/KEY: modified_base
(B) LOCATION:196
(D) OTHER INFORMATION:/mod_base= OTHER
/note= "Dideoxy T"

GCCATTGGTG	CTGTGGATCA	TGCGGGCGGC	CTGGGACTCC	TCCTGCTGGC	AGCGGGTCAG	60
CGACAGTGCG	TCGTCCATGT	GGCCCTCCTG	GTGCAGCATG	GCCTTCTTCT	TGAGCACGCC	120
GAGCCGCAGG	GCCTTGGGGT	CTGGAGCGGC	ATAGGTGAGC	CACAGTCCTG	AGGCCACATG	180
CTGCACGAAG	CACAGT					196

GCCATTGGTG	CTGTGGATCA	TGCGGGCGGC	CTGGGACTCC	TCCTGCTGGC	AGCGGGTCAG	60
CGACAGTGCG	TCGTCCATGT	GGCCCTCCTG	GTGCAGCATG	GCCTTCTTCT	TGAGCACGCC	120
GAGCCGCAGG	GCCTTGGGGT	CTGGAGCGGC	ATAGGTGAGC	CACAGTCCTG	AGGCCACATG	180
CTGCACGAAG	CACAGTGA	CT				200

(A) LENGTH: 204 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

(A) NAME/KEY: modified_base

(B) LOCATION: 204

(D) OTHER INFORMATION: /note= "Dideoxy T"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

GCCATTGGTG	CTGTGGATCA	TGCGGGCGGC	CTGGGACTCC	TCCTGCTGGC	AGCGGGTCAG	60
CGACAGTGCG	TCGTCCATGT	GGCCCTCCTG	GTGCAGCATG	GCCTTCTTCT	TGAGCACGCC	120
GAGCCGCAGG	GCCTTGGGGT	CTGGAGCGGC	ATAGGTGAGC	CACAGTCCTG	AGGCCACATG	180
CTGCACGAAG	CACAGTGA	CT				204

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 206 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "DNA generated by glycosylase mediated cleavage and followed by upstream fragment extension, and has a 3' hydrogen atom"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

(A) NAME/KEY: modified_base

(B) LOCATION: 206

(D) OTHER INFORMATION: /note= "Dideoxy T"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

GCCATTGGTG CTGTGGATCA TCGGGGCGGC CTGGGACTCC TCCTGCTGGC AGCGGGTCAG 60
CGACAGTGCG TCGTCCATGT GGCCCTCCTG GTGCAGCATG GCCTTCTTCT TGAGCACGCC 120
GAGCCGCAGG GCCTTGGGGT CTGGAGCGGC ATAGGTGAGC CACAGTCCTG AGGCCACATG 180
CTGCACGAAG CACAGTGACT CCCCCT 206

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 209 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "DNA generated by glycosylase mediated cleavage followed by upstream fragment extension, and has a 3' hydrogen atom"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

- (A) NAME/KEY: modified_base
- (B) LOCATION: 209
- (D) OTHER INFORMATION: /mod_base= OTHER
/note= "Dideoxy T"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

GCCATTGGTG CTGTGGATCA TCGGGGCGGC CTGGGACTCC TCCTGCTGGC AGCGGGTCAG 60
CGACAGTGCG TCGTCCATGT GGCCCTCCTG GTGCAGCATG GCCTTCTTCT TGAGCACGCC 120
GAGCCGCAGG GCCTTGGGGT CTGGAGCGGC ATAGGTGAGC CACAGTCCTG AGGCCACATG 180
CTGCACGAAG CACAGTGACT CCCCCTACT 209

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 204 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

000207 " 6EZE2960

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "DNA generated by glycosylase mediated cleavage followed by upstream fragment extension, and has a 3' hydrogen atom"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

(A) NAME/KEY: modified_base

(B) LOCATION:204

(D) OTHER INFORMATION:/mod_base= OTHER
/note= "Dideoxy C"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

GCCATTGCTG CTGTGGATCA TCGGGCGGC CTGGGACTCC TCCTGCTGGC AGCGGGTCAG	60
CGACAGTGGC TCGTCCATGT GGCCCTCCTG GTGCAGCATG GCCTTCTTCT TGAGCACGCC	120
GAGCCGCAGG GCCTTGGGGT CTGGAGCGGC ATAGGTGAGC CACAGTCCTG AGGCCACATG	180
CTGCACGAAG CACAGTGACT CCCC	204

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 54 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Homo sapiens

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

AACTTGTGGT AGTTGGAGCT GGTGGCGTAG GCAAGAGTGC CTTGACGATA CAGC

54

000201" 6E/E/960

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 54 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Generated by PCR amplification of genomic DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

AACTTGTGGT AGTTGGAGCT GGUGGCGUAG GCAAGAGUGC CUUGACGAUA CAGC

54

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 54 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Generated by PCR amplification of genomic DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

GCTGTATCGT CAAGGCACTC TTGCCTACGC CACCAGCUCC AACUACCACA AGUU

54

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 54 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Generated by PCR amplification of genomic DNA"

000201" 6622960

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

AACTTGTGGT AGTTGGAGCT GAUGGCGUAG GCAAGAGUGC CUUGACGAUA CAGC

54

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 54 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Generated by PCR amplification"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

GCTGTATCGT CAAGGCACTC TTGCCTACGC CAUCAGCUCC AACUACCACA AGUU

54

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 37 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Generated by glycosylase mediated cleavage of PCR amplified DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

GCTGTATCGT CAAGGCACTC TTGCCTACGC CACCAGC

37

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 32 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "Generated by glycosylase mediated cleavage of PCR amplified DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

GCTGTATCGT CAAGGCACTC TTGCCTACGC CA

32

(2) INFORMATION FOR SEQ ID NO: 22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 66 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

GCTGTAAACG ACGGCCAGTT TCATGCAGGG CTGGAGTCGT AGGCAAGAGT GCCTTGACGA
TACAGC

60

66

(2) INFORMATION FOR SEQ ID NO: 23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

GCTGTAAACG ACGGCCAGTT TCAT

24

000201" 6E/E/960

(2) INFORMATION FOR SEQ ID NO: 24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 66 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "Generated by primer extension"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

GCTGTATCGT CAAGGCACTC TTGCCTACGC CACCAGCCCT GCATGAACT GGCCGTCGTT 60
TACAGC 66

(2) INFORMATION FOR SEQ ID NO: 25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 66 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

GCTGTAAACG ACGGCCAGTT TCATGCAGGA TCCATGGCGT AGGCAAGAGT GCCTTGACGA 60
TACAGC 66

(2) INFORMATION FOR SEQ ID NO: 26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 66 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "Generated by primer extension"

000207" 6E2E2960

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

GCTGTATCGT CAAGGCACTC TTGCCTACGC CATGGATCCT GCATGAACT GGCCGTCGTT 60
TACAGC 66

(2) INFORMATION FOR SEQ ID NO: 27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

GGTAGTTGGA GCTGGTGGCG 20

(2) INFORMATION FOR SEQ ID NO: 28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

TCCAACCTACC 10

(2) INFORMATION FOR SEQ ID NO: 29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 47 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

000207" 6E/E/960

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "Generated by ligation of two DNA molecules"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

GCTGTATCGT CAAGGCACTC TTGCCTACGC CACCAGCTCC AACTACC

47

(2) INFORMATION FOR SEQ ID NO: 30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

CCAGCTCCAA

10

(2) INFORMATION FOR SEQ ID NO: 31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

TTGGAGCTGG TGGCGTAGGC

20

(2) INFORMATION FOR SEQ ID NO: 32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 42 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

000201" 6E/E/960

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "Generated by ligation of
two DNA molecules"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:

GCTGTATCGT CAAGGCACTC TTGCCTACGC CACCAGCTCC AA

42

000201" 6E/E/960